

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

SEP 26 2001

PATENT & TRADEMARK OFFICE

LAUSELA et al  
091,579,894  
0933-0159  
Birch, Stewart, Kofach  
& Birch LLP  
(703)205-8000

PA 241771

# THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

May 08, 2000

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: 60/136,085

FILING DATE: May 26, 1999



By Authority of the  
COMMISSIONER OF PATENTS AND TRADEMARKS

*W. Montgomery*  
W. MONTGOMERY  
Certifying Officer

SEP 28 2001

RECEIVED

1c604 U.S. PTO  
05/26/99

Approved

Please type a plus sign (+) inside this box → ☒

PTO/SB/16 (2-98)  
Approved for use through 01/31/2001. OMB 0651-0037  
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE  
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

+

# PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53 (c).

1c541 U.S. PTO  
60/136085  
05/26/99

INVENTOR(S)					
Given Name (first and middle [if any])		Family Name or Surname		Residence (City and either State or Foreign Country)	
KALLE MARTTI		SAKSELA		BROOKLYN, NY	
<input type="checkbox"/> Additional inventors are being named on the _____ separately numbered sheets attached hereto					
TITLE OF THE INVENTION (280 characters max)					
METHODS AND MATERIALS FOR GENERATING SH3 DOMAINS WITH TAILORED BINDING PROPERTIES					
Direct all correspondence to:			CORRESPONDENCE ADDRESS		
<input type="checkbox"/> Customer Number			Place Customer Number Bar Code Label here		
OR			Type Customer Number here		
<input checked="" type="checkbox"/> Firm or Individual Name			KALLE SAKSELA		
Address			INSTITUTE OF MED. TECHNOLOGY, UNIV. OF TAMPERE		
Address			POB 607		
City			State	ZIP	33101
TAMPERE					
Country			Telephone	Fax	
FINLAND			+358-3-2157029	+358-3-2158597	
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/> Specification Number of Pages		17		<input checked="" type="checkbox"/> Small Entity Statement	
<input checked="" type="checkbox"/> Drawing(s) Number of Sheets		3		<input type="checkbox"/> Other (specify) _____	
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)					
<input checked="" type="checkbox"/> A check or money order is enclosed to cover the filing fees				FILING FEE AMOUNT (\$)	
<input type="checkbox"/> The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: _____				75.00	
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.					
<input checked="" type="checkbox"/> No.					
<input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are: _____					

Respectfully submitted,

Date 5/20/99

SIGNATURE \_\_\_\_\_

REGISTRATION NO. \_\_\_\_\_

TYPED or PRINTED NAME KALLE SAKSELA

(if appropriate)

TELEPHONE +358-3-215 7029

Docket Number: \_\_\_\_\_

## USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is used by the public to file (and by the PTO to process) a provisional application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the complete provisional application to the PTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, Washington, D.C., 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, D.C., 20231.

+

<b>STATEMENT CLAIMING SMALL ENTITY STATUS (37 CFR 1.9(f) &amp; 1.27(b))--INDEPENDENT INVENTOR</b>	Docket Number (Optional)
---	--------------------------

Applicant, Patentee, or Identifier: KALLE MARTTI SAKSELA

Application or Patent No.: \_\_\_\_\_

Filed or Issued: \_\_\_\_\_

Title: METHODS AND MATERIALS FOR GENERATING SH3 DOMAINS WITH TAILORED  
BINDING PROPERTIES

As a below named inventor, I hereby state that I qualify as an independent inventor as defined in 37 CFR 1.9(c) for purposes of paying reduced fees to the Patent and Trademark Office described in:

- ☐ the specification filed herewith with title as listed above.
- ☒ the application identified above.
- ☐ the patent identified above.


I have not assigned, granted, conveyed, or licensed, and am under no obligation under contract or law to assign, grant, convey, or license, any rights in the invention to any person who would not qualify as an independent inventor under 37 CFR 1.9(c) if that person had made the invention, or to any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e).

Each person, concern, or organization to which I have assigned, granted, conveyed, or licensed or am under an obligation under contract or law to assign, grant, convey, or license any rights in the invention is listed below:

- ☒ No such person, concern, or organization exists.
- ☐ Each such person, concern, or organization is listed below.

Separate statements are required from each named person, concern, or organization having rights to the invention stating their status as small entities. (37 CFR 1.27)

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

<u>KALLE SAKSELA</u> NAME OF INVENTOR   Signature of inventor  <u>5/20/99</u> Date	NAME OF INVENTOR   Signature of inventor   Date	NAME OF INVENTOR   Signature of inventor   Date
--	---	---

669250" 98092799

# SPECIFICATION (17 PAGES)

## DESCRIPTION OF THE INVENTION:

### *Methods and materials for generating SH3 domains with tailored binding properties*

#### **Abstract:**

A number of biological processes that are important for normal and pathological states of are governed by interactions of cellular proteins mediated by Src homology region 3 (SH3) domains. This disclosure concerns methods and materials for generating SH3 domains with engineered binding properties, and their use as tools in research, therapy, and drug discovery.

#### **Background of the invention:**

Virtually all aspects of cellular behavior, such as adaptation of a cell in response to extracellular stimuli by changing of its pattern of gene expression, are regulated and executed by dynamic and ordered proximity of cellular proteins. During the evolution several different types of protein domains specialized in mediated such regulated and specific protein-protein interaction events have emerged. Protein domains of one type typically form large families of homologous but sufficiently divergent member, such that each of them have unique, although often overlapping specificities for ligand binding.

The SH3 domain was first identified as a region of homology among the Src family tyrosine kinases encoded by oncogenic retroviruses and their cellular proto-oncogene counterparts. Thereafter SH3 domains have been noticed in a large number (>50) of proteins that serve important functions in regulating cell growth, differentiation, and other processes. Because of these functions SH3 domains are intimately involved in pathogenesis of various diseases, in particular cancer. In addition, various microbial pathogens, such as HIV, exploit SH3-mediated processes as a part of their life-cycle. Ability to influence protein complex formation mediated by SH3 domains would therefore have significant therapeutic potential.

SH3 domains are small (typically 55-70 amino acids), modular protein units, and mediate inter- and intramolecular protein interactions by binding to ligands that contain a region with a secondary structure known as the polyproline type II (PPII) helix. Sequence variation in the PPII helix region of the ligand has been shown to influence the specificity in SH3/ligand complex formation. However, significant additional specific affinity is provided for different SH3 domain / ligand pairs by molecular contacts outside the SH3/PPII interface. An SH3 structure known as the RT-loop, which is highly divergent among the otherwise similar SH3 domains, has been shown to be involved in such specificity-determining interactions.

Previous attempts to generate molecules that could compete with naturally occurring SH3-interactions have focused on design or selection from random libraries of peptides and peptide-like molecules that could compete with PPII ligands for their binding to their cognate SH3 domains. Success in such approaches has been reported by a number of groups (see references in "Example1"), and patent application for different modification of this approach have been filed (such as PCT WO95/24419 and PCT WO96/03649).

However, the relative similarity of the SH3/PPII interface of different SH3/ligand pairs presents a problem for developing highly specific inhibitory molecules. To overcome this problem we have chosen a different approach, which is based on the apparent role of the SH3 domain RT-loop in ligand selection that has been indicated by a number of studies, in particular our previous work on the complex between the HIV Nef protein and the SH3 domain of the cellular Hck tyrosine kinase. As explained below in, we have found that by randomly manipulating the amino acid sequence comprising the variable region of the RT-loop (in this case six amino acids of the Hck SH3 domain) it is possible to create artificial SH3 domains that bind with unnaturally high affinities and with predetermined binding specificities to different ligand proteins.

### Summary of the Invention:

The well-characterized interaction between HIV-1 Nef and the SH3 domain of Hck is one the tightest known SH3-mediated interactions. We have previously shown that a similar capacity for binding to Nef can be transferred to Fyn-SH3 by engineering Hck-like amino acid substitutions into its RT-loop. The present invention is that, instead of mimicking the structure of a naturally occurring, known cognate SH3 domain, one can generate SH3 domains with desired ligand binding properties by using random manipulation of the RT-loop sequence combined with a powerful functional selection. Notably the method described in this invention can be used to identify SH3

domains with unnaturally high affinities specific for proteins known to bind to any naturally occurring SH3 domains, as well as to target proteins that believed to be SH3 ligands but lack an identified SH3 domain-containing cellular partner.

In the prototypic version of the invention presented here we have used polymerase chain reaction (PCR) assisted mutagenesis to produce a large collection of SH3 domains (termed RRT-SH3) that shared their overall structure with Hck-SH3, but contained a random combination of amino acids in place of the six RT-loop residues that constitute a variable region among the otherwise homologous SH3 domains. We generated  $\approx 130$  million individual recombinant phagemid-expressed RRT-SH3 domains, which exceeds the theoretical number of all possible combinations of six residue peptides made of the 20 different amino acids ( $20^6 = 64$  million). We discovered that a method known as phage-display is well-suited for expression of functional SH3 domains on the surface of bacteriophage particles, and used this method as a selection tool for identification of the rare RRT-SH3 domains with desired binding properties from among the millions of non-binding molecules. Although at the moment phage-display appears to be the preferred method for such functional selection of the RRT-SH3 domains, the current invention is not restricted to this approach, and a number of other methods could also be used for this task.

As a proof of principle for the usefulness of this approach we showed that it is possible to identify individual RRT-SH3's which bound to HIV-1 Nef with up to 40-fold higher affinities than the already avidly binding (KD 250 nM) Hck-SH3 domain that served as the backbone of the artificial RRT-SH3 molecules. Such superior affinity provided these molecules an ability to efficiently compete with the Hck/Nef interaction even when present in low concentrations, thus pointing out to obvious therapeutic applications. Notably, we showed that it is possible to identify RRT-SH3 molecules from this Hck-derived library that bind with very high affinities to SH3 ligand protein that do not have significant affinity for the natural, unmodified Hck-SH3. As an example of such a protein we used a mutant version of the Nef protein (NefR90) whose ability to bind Hck-SH3 has been disturbed by a mutation that specifically prevents the accommodation of the native Hck-SH3 RT-loop. These proof-of-principle studies are described below in "EXAMPLE 1: SH3 domains targeted with high affinity to wild-type and Hck-SH3 binding-deficient forms of HIV-1 Nef", together with methodological details of this prototypic application of the invention as well as an up-to-date review of the literature in this field of research.

The present invention provides a generally useful method for targeting of SH3 ligand proteins with artificial SH3 domains. Although our results indicate that very different binding specificities can be generated by manipulation of the SH3 RT-loop of a single (Hck) SH3-domain backbone structure, other SH3 domains can be similarly modified by the described methodology, if necessary in order

to create high affinity molecules for all SH3 ligand proteins. Also, the engineering of the binding-properties by manipulation of the RT-loop region could be combined with other random or specific modifications of the SH3 domain to further optimize its binding or other physico-chemical properties, such as protease resistance or solubility.

Several valuable applications for tailor-made SH3 domains described by the present invention can be envisioned. They could introduced into cells in order to inhibit or otherwise regulate the functions SH3-ligand proteins and their complexes of with other proteins. If the RRT-SH3 domains would be equipped protein with an appropriate targeting signal the subcellular localization of their target protein could also be influenced this way. While the most apparent uses of such approaches are in basic research on SH3-mediated cellular processes, similar applications in gene therapy are also obvious. The relatively small size of an SH3 domain also makes it possible that the engineered RRT-SH3 proteins could be delivered into target cells by fusing them to membrane-penetrating peptides or other such methods. Besides being used as therapeutic agents themselves, structural analysis of the RRT-SH3/ligand complexes could also help in design of non-peptide compounds targeted against the SH3-binding interfaces of various disease-associated cellular proteins. Finally, the ability of RRT-SH3 to interact tightly to the SH3-binding surface of the ligand proteins, and thereby compete with any compounds that bind to the same region, could also be helpful in identifying molecules derived by non-rational drug discovery screens that are targeted against these functionally important regions.



## Brief description of the drawings:

**Figure 1.** A simplified cartoon depicting the strategy of current invention for creating SH3 domains with altered binding properties. Interactions of SH3 domains, such as Hck-SH3, with their cognate ligand proteins involve a set of molecular contacts between the SH3 domain and a proline repeat (PxxP) motif region in the ligand protein. Although necessary for binding, these PxxP-directed interactions are weak and relatively similar in most SH3/ligand complexes, and therefore provide little selectivity and strength for these interactions. Our previous studies on the Hck-SH3/HIV Nef complex have revealed that a distinct and divergent region in the SH3 domain (the RT-loop) can provide strong affinity and significant specificity for binding. Prompted by these observations we have generated a large collection of artificial SH3 domains in which the six amino acids of the Hck-SH3 RT-loop (EAIHHE) that provide its affinity for Nef have been replaced with a random hexapeptide, and shown that from this library of randomized molecules individual SH3 domains with desired binding properties can be identified.

**Figure 2.** Data from two representative experiments measuring binding of four different RRT-SH3 domains to immobilized Nef based on their ability to compete with labeled Hck-SH3 (66 nM). The degree of inhibition of probe binding to Nef caused by the added competitors is shown on the y-axis, and the 50% level is indicated by the horizontal dashed line. The concentration of the competitor relative to the probe (ranging 1:128 to 4:1 for Hck-SH3, and 1:8 to 256:1 for RRT.A1) is shown on the x-axis, and the point where the probe and the competitor were present in equimolar amounts is indicated by the vertical dashed line.

**Figure 3.** Binding of selected RRT-SH3 domains to a Nef variant carrying a F90R mutation. NefR90 was used a ligand for biotinylated RRT.A1 (66 nM) in the presence of unlabeled SH3 domains at relative concentrations ranging from 1:256 - 16:1. Summary of similar data on all tested RRT-SH3 domains is presented in Table I.

## EXAMPLE 1: SH3 domains targeted with high affinity to wild-type and Hck-SH3 binding-deficient forms of HIV-1 Nef

### Summary

The avid binding of HIV-1 Nef to the Src homology-3 (SH3) domain of Hck ( $K_D$  250 nM) has been shown to involve an interaction between the RT-loop of Hck-SH3 and residues in Nef outside of its prototypic polyproline type II (PPII) helix-containing SH3-ligand region. Such distinctive interactions are thought to provide specificity and affinity for other SH3/ligand protein complexes as well. In this study we have constructed and successfully displayed on the surface of M13 bacteriophage particles a complex library of SH3 domains, which are derived from Hck but carry a random hexapeptide substitution in their RT-loops (termed RRT-SH3). By this strategy we have identified individual RRT-SH3 domains that can bind to Nef with affinities higher than  $K_D$  10 nM. Some of these high-affinity RRT-SH3 domains resembled Hck-SH3 in that they bound much less well to a Nef variant containing an engineered F90R mutation that interferes with docking of the native Hck RT-loop. In addition, we could also select RRT-SH3 domains with an opposite specificity, which were dependent on the Arg<sup>90</sup> residue for strong binding, and had a 100-fold lower affinity for unmodified Nef. These results emphasize the importance of the RT-loop in SH3 ligand selection, and suggest a general strategy for creating SH3 domains with desired binding properties.

### Materials and methods

To produce a phagemid for expression of native Hck-SH3, a DNA fragment encoding for a 56 aa Hck polypeptide fragment was amplified by PCR using Pfu polymerase (Stratagene) and primers containing Pst-I and Not-I sites, and inserted into the corresponding sites in pCANTAB-5EP, a modified version of pCANTAB-5E (Pharmacia) with a new Pst I cloning site. To create the library of RRT-SH3 domains, a longer sense PCR primer 5'-AAT CTG CAG GAA TTC GTG GTT GCC CTG TAT GAT TAT NNN NNK NNS NNK NNK NNS GAC CTC AGC TTC CAG AAG GGG GAC-3' extending over the RT-loop encoding region of Hck was used, and the resulting fragment was cloned into pCANTAB-5EP as described above. A total of  $137 \times 10^6$  individual recombinant colonies were obtained by electroporation into E. coli TG1 cells, which were pooled and infected with the M13KO7 helper phage as explained below.

Overnight cultures of TG1 cells carrying the phagemid(s) of interest were grown at 30°C in 2xYT containing 100 ug/ml ampicillin and 2% glucose (2xYT/AG). Cultures were diluted 1/10 in 2xYT/AG, and incubated with  $5 \times 10^8$  pfu/ml of M13KO7 helper phage (Pharmacia) for 2 hours at 37°C. The bacteria were pelleted, and the medium changed to 2xYT containing 100 ug/ml ampicillin 50 ug/ml kanamycin (2xYT/AK). After overnight growth of the double resistant bacteria in this medium at 37°C, the supernatant containing the recombinant phages was collected, passed through a 0.45 mm filter, and stored in aliquots at 4°C until needed.

Six-well plates were coated with 10 ug/ml GST-Nef (HIV-1 NL4-3 Nef containing a T71R change to mimic Nef from typical primary HIV-1 isolates; Saksela *et al.*, 1995), GST-NefR90 (see text; Manninen *et al.* 1998. *Virology*, 250, 273-82), GST-Nef-PA1 (P72A+P75A mutant of Nef; Saksela *et al.* 1995. *EMBO J*, 14, 484-91) or plain GST in 50 mM sodium carbonate pH 9.6 at 4°C overnight. Nonspecific binding sites were blocked with 5% milk in PBS/0.05% Tween 20, and the wells washed briefly with PBS before  $10^7$ - $10^{11}$  pfu of recombinant phages per well were added (higher titers used in the early rounds of selection), followed by incubation for 2 hours at RT. In some experiments involving NefR90-coated wells, the phage solution was supplemented with 10 ug/ml of soluble wild-type Nef. After incubation with the phages the wells were washed six times (5 min) with PBS + 0.05% Tween 20 and three times with PBS. In most experiments the bound phages were eluted with a small volume of PBS containing an excess of (150 ug/ml) of the same Nef protein immobilized in the well. Alternatively, the TG1 cells to be infected were added directly to the washed wells. In both cases, the bacteria were first grown in 2xYT to log phase from an overnight culture, infected with the affinity-selected phages for 2 hours at 37°C, after which a sample of 1% was removed for determination of the infectious titer of the selected phages by plating on ampicillin plates. These plates also served as indicators for the enrichment of specific clones when compared to plates infected with phages from GST-coated wells processed in parallel. The remaining 99% of the infected bacteria were supplemented with 100 ug/ml ampicillin and 2% glucose, and subjected to a subsequent infection with  $5 \times 10^8$  pfu/ml of M13KO7 helper phages for 2 hours at 37°C, after which they were pelleted and resuspended into an equal volume of 2xYT/AK. After an overnight incubation the amplified recombinant phage supernatants were collected as described above, and used for the subsequent round of selection/infection. Usually after 8 cycles of selection  $\geq 12$  colonies were picked for miniprep preparation of phagemid DNA, and their RRT-SH3 insert sequenced using ABI Prism 310 (Perkin Elmer Applied Biosystems).

Generation of the pGEX vector for bacterial expression of the GST fusion proteins for Nef and Hck-SH3 has been described previously (Saksela *et al.* 1995. *EMBO J*, 14, 484-91; Manninen

et al. 1998. *Virology*, 250, 273-82). GST-RRT-SH3 vectors were constructed by PCR amplification of the SH3 fragments from the corresponding phagemids, and insertion between the Eco RI and Sal I sites in pGEX-4T-1 (Pharmacia), and verified by sequencing. Expression and purification of the GST fusion proteins in *E. coli* BL21 were carried out by standard methods as recommended by the supplier of the pGEX vectors and glutathione resin (Pharmacia). After elution of the fusion proteins from the resin they were concentrated and their buffer changed to PBS by ultrafiltration in Centrex UF2 columns (Shleicher & Schuell), followed by concentrations measurement using the BioRad (Lowry) method using BSA as a standard. The concentration and integrity of the proteins were further confirmed in Coomassie blue stained SDS PAGE gels. Biotinylation was done using the EZ-Link Sulfo-NHS-LC-Biotin reagent, as suggested by the manufacturer (Pierce), after which these proteins were subjected to 3 rounds of ultrafiltration to remove any free biotin, followed by concentration measurement as described above.

Nunc Maxisorb F8 strips were coated with the different GST-Nef proteins (200 ng in 100  $\mu$ l per well) overnight at 4°C, followed by a 30 min incubation at RT with 1.5% BSA in washing buffer (WB; PBS + 0.05% Tween-20) to saturate non-specific protein binding, and washed twice with WB. The unlabeled SH3 proteins used as competitors were diluted into WB that contained 1.5% BSA and a large molar excess of plain GST (4  $\mu$ M) to minimize any GST-directed dimerization of the fusion proteins. 50  $\mu$ l of this solution was mixed with an equal volume of the probe (biotinylated SH3 in WB) and added to the wells resulting in a final probe concentration of 66 nM (Hck-SH3/Nef and RRT.A1/NefR90 assays) or 7 nM (RRT.A1/Nef assay). After a 1 hour incubation at RT the wells were washed three times with WB, and added 100  $\mu$ l of 1:2000 dilution (in WB) of streptavidin-biotin horseradish peroxidase complex (Amersham Life Sciences) per well. The plates were incubated 45 minutes at RT and washed again three times, after which their peroxidase activity was measured using 1,2-phenyldiamine-dihydrochloride (OPD; 0.6 mg/ml; Fluka AB) as a substrate. The enzymatic reactions were stopped after 10 minutes by adding 50  $\mu$ l of 2 M sulfuric acid, followed by optical density measurement at 492 nm using a Victor 1420 Multilabel Counter (Wallac).

## Introduction

SH3 domains are globular protein modules typically comprising of 50 - 60 amino acids found in many different proteins, particularly proteins involved in cellular signal transduction (Cohen et al. 1995. *Cell*, 80, 237-48; Dalgarno et al. 1997. *Biopolymers*, 43, 383-400). SH3 domains mediate inter- and intramolecular interactions by binding to ligands that contain a region with a secondary structure known as the polyproline type II (PPII) helix. These ligands can bind to SH3

domains in two opposite orientations, and typically show the "PxxP motif" consensus sequences RXØPXXP and PXØPXR (Ø is a hydrophobic amino acid, X is any amino acid) (Feng et al. 1994. *Science*, 266, 1241-7; Lim et al. 1994. *Nature*, 372, 375-9). The positioning of the conserved basic residue (usually an arginine; R) in the PxxP motif determines in which orientation the ligand binds its cognate SH3 domain. In addition, there are atypical SH3-ligands with PPII helices that do not conform to such consensus rules. A notable example is the PPII region in Src, which is involved in catalytic autoinhibition by binding to the SH3 domain of Src itself, but contains only one of the two prolines that ordinarily define a PxxP-motif (Xu et al. 1997. *Nature*, 385, 595-602).

Sequence variation in the PPII helix region, involving the consensus as well as the adjacent, non-consensus positions, has been shown to influence the specificity in SH3/ligand complex formation. Examples of preference for targets with atypical PxxP consensus motifs have been provided by studies addressing Abl SH3 ligand selection (Feng et al. 1994. *Science*, 266, 1241-7; Weng et al. 1995. *Mol Cell Biol*, 15, 5627-34; Yu et al. 1994. *Cell*, 76, 933-45), and the CrkN-SH3/C3G peptide complex (Knudsen et al. 1995. *EMBO J*, 14, 2191-8; Wu et al. 1995. *Structure*, 3, 215-26). The effect of sequence variation involving the non-consensus residues in the PPII region of SH3-ligands has been best demonstrated by experiments in which distinctive target sequences have been selected for different SH3 domains from libraries of chemically synthesized or phage-displayed random peptides (Sparks et al. 1994. *J Biol Chem*, 269, 23853-6; Viguera et al. 1994. *Biochemistry*, 33, 10925-33; Yu et al. 1994. *Cell*, 76, 933-45). However, despite the above-discussed evidence for specificity, the maximal SH3-binding affinities of short PPII ligand peptides are low, and the relative differences in their binding to different SH3 domains are modest.

By contrast, there is increasing evidence that molecular contacts outside the PPII helix interface can provide significant specificity and strength to SH3-binding. Use of phage-display libraries of longer peptides containing a PxxP motif embedded within random sequence has demonstrated that the flanking residues can increase the selectivity of such ligands, which may show up to 20-to-30 fold differences in their affinities towards different SH3 domains (Rickles et al. 1994. *EMBO J*, 13, 5598-604; Rickle et al. 1995. *Proc Natl Acad Sci U S A*, 92, 10909-13; Sparks et al. 1996. *Proc Natl Acad Sci U S A*, 93, 1540-4). Structural analysis of the interactions of Src-SH3 with two such dodecapeptides revealed that the relatively high specificity and affinity ( $K_D$  values 0.54  $\mu$ M and 1.2  $\mu$ M) of these interactions involved contacts between the flanking residues in the peptides and two loop-like structures in the Src-SH3 domain, which represent regions of high sequence diversity among different SH3 domains and are known as the n-src- and RT-loops (Feng et al. 1995. *Proc Natl Acad Sci U S A*, 92, 12408-15). Similarly, the specific binding of a rationally designed proline-rich ligand to Abl SH3 ( $K_D$  0.4  $\mu$ M for Abl vs. 273  $\mu$ M for

Fyn-SH3) could be explained by corresponding molecular contacts with Abl SH3 (Pisabarro and Serrano. 1996. *Biochemistry*, 35, 10634-40; Pisabarro et al. 1998. *J Mol Biol*, 281, 513-21).

Another interaction that has been informative in elucidating the basis of SH3 binding specificity, which also emphasizes the role of the RT-loop, is the complex between HIV-1 Nef and the SH3 domain of the tyrosine kinase Hck. Nef is a 27 - 34 kD myristoylated protein of primate lentiviruses (HIV-1, -2, and SIVs), and important for development of high viremia and immunodeficiency in the infected host Harris. 1996. *J Gen Virol*, 77, 2379-92; Saksela. 1997. *Front Biosci*, 2, 606-18). Interestingly, Nef has remarkably selective SH3-binding characteristics. It can bind tightly to the Hck-SH3, showing affinity values of approximately  $K_D$  0.2  $\mu$ M as measured by surface plasmon resonance (Lee et al. 1995. *EMBO J*, 14, 5006-15). In contrast to the strong binding to Hck, Nef has almost a 100-fold lower affinity towards the highly homologous SH3 domain of Fyn. Biochemical and structural studies have revealed that the basis of this selectivity lies in the efficient strategy of Nef for recognition of the non-conserved SH3 residues distinctive to Hck, in particular the side chain of an isoleucine located in the RT-loop of Hck-SH3 (Lee et al. 1996. *Cell*, 85, 931-42). The region that accommodates the Hck-SH3 RT-loop is composed of multiple non-contiguous parts of the Nef polypeptide, and is located distally from the PPII region in the three-dimensional structure of Nef.

Taken together, the above observations suggest a general model where regions in SH3 ligands outside the PPII helix region provide specificity and affinity for binding by contacting regions that are divergent among SH3 domains, in particular residues in the RT-loop. Prompted by this concept, in this study we have constructed a large library (>130 millions) of Hck-derived artificial SH3 domains, in which six non-conserved, Hck-specific residues in the RT-loop have been replaced by a random hexapeptide (termed RRT-SH3 for randomized RT-loop), and expressed these on the surface M13 bacteriophages in order to identify novel SH3 domains with engineered binding characteristics. We show that phage-display is well suited for presentation and selection of modified SH3 domains, and provide strong experimental support for a role of the SH3 RT-loop as a versatile specificity and affinity determinant.

## Results

To examine if it would be possible to present functional SH3-domains on the surface of bacteriophages we constructed a phagemid containing Hck-SH3, and by using a M13KO7 helper virus produced recombinant phages expressing it fused to the pIII coat protein. When these phages were incubated in six-well plates coated with purified GST-Nef, GST-Nef-PA1 (an SH3 binding deficient P72A;P75A mutant of Nef), or plain GST protein, the titers of infectious phage-particles

resisting washes were found to be 2 to 3 orders of magnitude higher in the GST-Nef coated wells as compared to wells coated with plain GST or with the GST-Nef-PA1 fusion protein (data not shown). Thus, these results indicated that functional SH3 domains can be displayed on the surface of M13 phage particles, and that such phages can be positively selected using a cognate SH3 ligand as an affinity reagent.

To create SH3 domains with novel ligand binding properties, we used PCR-assisted mutagenesis to produce a large collection of SH3 domains (termed RRT-SH3) that shared their overall structure with Hck-SH3, but contained a random combination of amino acids in place of the six RT-loop residues that constitute a variable region among the otherwise homologous SH3 domains. We generated  $\approx 130$  million individual recombinant phagemid clones expressing different RRT-SH3 domains, which exceeds the theoretical number of all possible combinations of six residue peptides made of 20 different amino acids ( $20^6 = 64$  million). An infectious phage stock displaying this heterogeneous population of RRT-SH3 proteins was generated, as explained above for the Hck-SH3-expressing phagemid, and used for subsequent affinity selection experiments.

To test if the RRT-SH3 library contained novel Nef-binding SH3 proteins, we carried out an extensive selection of phages showing affinity for wild-type HIV-1 Nef. The selection scheme consisted of multiple (11 in all) independent experiments consisting of 7 or 8 rounds of affinity-selection and helper virus-assisted reamplification of the recombinant phages. A representative collection of RRT-SH3 clones derived from different rounds of phage selection were sequenced to follow the loss of heterogeneity in the population and the emergence of putative Nef-binding RRT-SH3 clones. A total of 278 RRT-SH3 clones selected using Nef were sequenced (Table I and data not shown). After six rounds of selection, the phage populations were typically dominated by 1 - 3 different RRT-SH3 clones. In most cases these dominant clones were identical or similar to the clones derived from other independent experiments, leading us to believe that we had rather exhaustively sampled the library for SH3 domains capable of high-affinity binding to HIV-1 Nef.

The RT-loop sequences of the RRT-SH3 clones found to be dominant after  $\geq 7$  rounds of selection are listed in Table I. They fell into four sequence families characterized by three or more identical residues within each class. In addition, these families also shared obvious similarities, most notably a serine residue in the second of the six (2/6) randomized positions present in all of them. One prevalent clone (RRT.C1) contained the consensus motifs of two different RT-loop sequences families (B and C). Besides the frequent appearance of serine, these RT-loops were remarkably rich in aromatic and proline residues. Although the RT-loop of the natural Nef ligand, Hck-SH3, also contains two aromatic residues (histidines) the selected RT-loop sequence did not

bear obvious resemblance to this region in Hck-SH3, or any other naturally occurring SH3 domain in the GenBank database.

SH3 domain (Competitor)	RT-loop	Probe: Hck-SH3	Probe: RRT.A1	
		Comp./probe at 50% inhib. Ligand: Nef	Inhib. with 1:1 comp./probe ratio Ligand: Nef	Ligand: NefR90
Hck-SH3	E A I H H E	1: 1	<5 %	<5 %
<b>Nef-selected:</b>				
RRT.A1	V S W S P D	1: 35	55 %	52 %
RRT.A2	F S W S D T	1: 19	20 %	53 %
RRT.A3	D S W S T S			
RRT.A4	Y S W S D M	1: 10	34 %	44 %
RRT.B1	W S P F P S	1: 26	41 %	9 %
RRT.B2	D S P F S F	1: 22	22 %	26 %
RRT.B3	F S P F S F			
RRT.B4	F S P F D W	1: 21	41 %	37 %
RRT.B5	S S P F D W			
RRT.B6	Y S P F S W	1: 37	51 %	14 %
RRT.C1	T S P F P W	1: 31	39 %	<5 %
RRT.C2	Y S F F P W	1: 16	20 %	13 %
RRT.C3	Y S D F P W	1: 26	42 %	20 %
RRT.C4	D S W F P W	1: 14	10 %	19 %
RRT.D1	S S F Y S S	1: 22	45 %	17 %
<b>NefR90-selected:</b>				
RRT.m1	Q G F L D Q	1: 0.8	<5 %	73 %
RRT.m2	N A F L P S	1: 2	<5 %	73 %
RRT.m3	E A W S P L	1: 17	22 %	62 %
RRT.m4	E S Y S E W	1: 0.8	<5 %	46 %

**Table I.** Phage-selected RT-loop sequences and binding properties of the corresponding SH3 domains

When the RRT-SH3 phage supernatants from the late rounds of selection, which were dominated by the clones shown in Table I, were subjected to an additional round of affinity selection using wild-type Nef and the SH3-binding-negative Nef-PA1 mutant in parallel, efficient capture of recombinant phages was seen only in the former case, whereas phage recovery from the



Nef-PA1-coated wells did not differ from the control wells without Nef (data not shown). Thus, while the RT-loop sequences of the enriched RRT-SH3 phages provided them with a superior affinity for Nef as compared to the rest of the clones in the library, these interactions were critically dependent on an intact PxxP-motif in Nef, as expected for *bona fide* SH3-mediated binding.

To study the binding properties of the selected RRT-SH3 domains in more detail, we transferred several members from each of the clone families into a bacterial expression vector to be produced for biochemical studies. To estimate their affinities for Nef we developed a competitive 96-well plate binding assay. In this assay serial dilutions of RRT-SH3 proteins were tested for their ability to compete with a constant amount of biotinylated Hck-SH3 for binding to immobilized Nef, followed by colorimetric quantitation of bound Hck-SH3. The binding affinities of the different RRT-SH3 proteins used as competitors could then be easily calculated based on the well-characterized  $K_D$  value (250 nM) of the Hck-SH3/Nef interaction. This assay design was chosen because it should be relatively insensitive to most potential sources of error that might easily affect a non-competitive assay, such as effects caused by the GST part of the SH3 fusion proteins or related to the incubation and washing conditions. The validity of this assay was strongly supported the excellent correlation of the theoretical and experimental values for a homotypic competition of unlabeled Hck-SH3 with biotinylated Hck-SH3, showing close to 50% inhibition at 1:1 ratio, and expected values over a wide range of concentration ratios (Fig. 2).

As evident from Figure 2 and Table I, all RRT-SH3 proteins from late rounds of selection bound to Nef substantially better than Hck-SH3. Many of these RRT-SH3's were able to compete more than 50% of binding of Hck-SH3 to Nef even when present in a 20-fold lower concentration, indicating a  $K_D$  value below 12 nM. The proteins RRT.A1, RRT.B6, and RRT.C1 could consistently inhibit 50% or more of the Hck-SH3 binding even at concentration ratios lower than 1:35, corresponding to  $K_D$  values as low as 7 nM.

As discussed above, an isoleucine residue in the RT-loop of Hck-SH3 has been shown to be important for its binding to Nef. The aliphatic side chain of this Ile residue fits into a hydrophobic pocket formed in part by the Nef residue Phe<sup>90</sup>. Mutation of this phenylalanine into an arginine (creating a mutant referred to as NefR90 in the following) greatly diminishes the affinity of Hck-SH3 binding ( $K_D$  1.99  $\mu$ M). While this interaction provides a structural explanation for the affinity of Hck-SH3, other modes of molecular recognition of the Nef SH3-ligand surface could also be envisioned. Therefore, we were interested in examining whether the Nef-selected RRT-SH3 domains also depended on a similar hydrophobic interaction, and would therefore also be affected by the F90R mutation in Nef.

Due to the low affinity of the Hck-SH3/NefR90 interaction it was not possible develop a quantitative 96-well plate assay based on binding of Hck-SH3 to immobilized NefR90. In contrast, strong binding and an adequate signal-to-noise ratio were observed when biotinylated RRT.A1 protein was tested as a probe. This indicated that the RRT-SH3 domain RRT.A1 did not critically depend on the Phe<sup>90</sup> residue in its binding to Nef, and allowed us to examine the relative affinities of the rest of the RRT-SH3 proteins for NefR90 based on their ability to compete with RRT.A1 (Fig. 3). As expected, Hck-SH3 was a very inefficient competitor of the RRT.A1/NefR90 interaction (Fig. 3 and Table I), and failed to cause a 50% inhibition of binding even when used in a 16-fold molar excess (66 nM vs. 1.1 mM), which was the highest practical competitor:probe ratio in the NefR90-assay. Based on the  $K_D$  value 1.99 mM previously measured for the Hck-SH3/NefR90 complex (Manninen et al. 1998. *Virology*, 250, 273-82), the absolute affinity of RRT.A1 for NefR90 was therefore estimated to be 100 nM or better. When the other Nef-selected RRT-SH3 proteins were tested as competitors for RRT.A1 binding, widely varying affinities for NefR90 were observed (Fig. 3 and Table I). Some RRT-SH3's (such as RRT.A2 and RRT.B4) competed efficiently with RRT.A1 binding, indicating that like RRT.A1 they also were relatively insensitive to the F90R change. In contrast, other RRT-SH3's (such as RRT.B1 and RRT.C1) were almost completely unable to interfere with RRT.A1/NefR90 complex formation even when tested at high concentrations. Thus, some of the high-affinity RRT-SH3 proteins shared with Hck-SH3 a mode of binding to Nef in which the Nef Phe<sup>90</sup> plays a dominant role, whereas others appeared to recognize Nef by dissimilar mechanisms, and bound well also to NefR90.

The above finding suggested that, depending on their RT-loops, the selected SH3 domains could recognize ligands by divergent molecular strategies. Encouraged by this observation we wanted to extend this concept further by engineering SH3 domains with entirely redirected binding specificities. To this end, we decided to develop RRT-SH3 domains that would not bind to wild-type Nef, but instead would show strong affinity for the NefR90 mutant. Thus, binding of such RRT-SH3 domains would be strictly dependent on interactions mediated by structural determinants involving the mutant Arg<sup>90</sup> residue, and would bear little resemblance to the natural recognition of Nef by Hck-SH3. For this purpose Nef-F90R was used as an affinity selection reagent, as described above for wild-type Nef, except that in some experiments an excess of soluble wild-type Nef was also included to enhance selection RRT-SH3 proteins that would primarily recognize Arg<sup>90</sup>-dependent determinants.

The most frequently observed RT-loop sequences of the NefR90-selected RRT-SH3 domains are shown in the bottom panel of Table I. Although these selection experiments were carried out almost as extensively (6 independent experiments) as in the case of wild-type Nef, equally obvious consensus RT-loop sequences were not apparent. As seen in Table I, however,

certain amino acids were favored in some positions, such as the dipeptide Phe-Leu in the positions 3/6 and 4/6, Glu in 1/6, Ala in 2/6, and Ser in 4/6. Notably, these NefR90-selected RT-loops did not contain any of the sequence motifs that characterized the Nef-selected RRT-SH3 domains. Also, with one exception (RRT.m4), the Ser 2/6 that was invariant in the Nef-selected clones was absent from the NefR90-selected RRT-SH3 domains.

As shown in Fig. 3 and Table I, all RRT-SH3 proteins selected for NefR90 binding could efficiently compete with binding of the clone RRT.A1 to NefR90, and were estimated to have affinities 8- to 4-fold higher than the "dual-specific" clone RRT.A1. Remarkably, binding of some of these NefR90-selected RRT-SH3 domains was critically dependent on the mutant Arg<sup>90</sup> residue, whereas others appeared to be "dual-specific". An example of the latter was RRT.m3, which besides being a potent inhibitor of the RRT.A1/NefR90 interaction, also competed well with RRT.A1 in binding to Nef. In contrast, the clones RRT.m1 and RRT.m4 were virtually unable to compete with RRT.A1 for binding to Nef (Fig. 3 and Table I). Thus, only a small fraction of the strong affinity of RRT.m1 and RRT.m4 towards NefR90 was contributed by structural determinants that did not involve the mutant Arg<sup>90</sup> residue. Due to this residual (Arg<sup>90</sup>-independent) affinity, however, they could still compete with the Hck-SH3/Nef interaction. Nevertheless, the fact these clones bound to NefR90 at least 100-times better than Hck-SH3 or some of the Nef-selected clones (such as RRT.C1) clearly indicated that they had been provided with a strikingly novel binding specificity, which targeted them for a high-affinity interaction that depended on an artificially created epitope in their ligand.

## Discussion

The well-characterized interaction between HIV-1 Nef and the SH3 domain of Hck is one of the tightest known SH3-mediated interactions. We have previously shown that a similar capacity for binding to Nef can be transferred to Fyn-SH3 by engineering Hck-like amino acid substitutions into its RT-loop (Lee et al. 1995. EMBO J, 14, 5006-15). The present study was undertaken to examine if, instead of mimicking the structure of a known cognate SH3 domain, it would be possible to generate SH3 domains with desired ligand binding properties by using random manipulation of the RT-loop sequence combined with a powerful functional selection. The data presented here show that this approach is indeed feasible, and suggest that it could be generally useful in elucidation of the structural basis of protein-protein interactions, and more specifically, in studies on biology and therapeutic applications of cellular processes that involve SH3-mediated protein interactions.

By displaying a library of SH3 domains carrying a random six amino acid substitution in their RT-loops (RRT-SH3 domains), we could select individual RRT-SH3's which bound to HIV-1 Nef

with up to 40-fold higher affinities than the avidly binding parental Hck-SH3 domain. Interestingly, these RRT-SH3 molecules recognized Nef by dissimilar molecular strategies, as indicated by large differences in the sensitivity of their binding to a F90R mutation in Nef (NefR90). This mutation has previously been shown to decrease the affinity of Hck-SH3 binding by preventing efficient docking of its RT-loop (Lee et al. 1996. *Cell*, 85, 931-42; Manninen et al. 1998. *Virology*, 250, 273-82). While some of the RRT-SH3 domains selected with wild-type Nef also bound well to NefR90, others were dependent on the presence of the Nef Phe<sup>90</sup> residue even more critically than Hck-SH3, and showed over a 100-fold decreased affinity towards NefR90. Conversely, by screening the RRT-SH3 library using NefR90, we were able to generate SH3 domains whose high-affinity binding depended on structural determinants involving the side-chain of the mutant Arg<sup>90</sup> residue. Thus, depending on the hexapeptide engineered into the RT-loop of a single SH3 domain (Hck), a dramatic increase in the affinity for a known ligand (Nef), or alternatively high-affinity binding to a novel ligand (NefR90), was observed. Understanding the structural basis of molecular recognition and generation of binding affinity in these interactions poses an interesting challenge for future work, and could provide novel insights into the principles of protein-protein communication.

Several reports indicate that an important role of the RT-loop in SH3 ligand selection is probably universal, rather than specific for Nef/SH3 complexes (Dalgarno et al. 1997. *Biopolymers*, 43, 383-400; Feng et al. 1995. *Proc Natl Acad Sci U S A*, 92, 12408-15; Feng et al. 1996. *Chem Biol*, 3, 661-70; Pisabarro et al. 1998. *J Mol Biol*, 281, 513-21). It is therefore likely that avidly binding SH3 domains specific for many proteins, which normally serve as ligands for other or yet unknown SH3 domains, could be selected from our Hck-derived RRT-SH3 library. However, it is possible that due to restrictions imposed by the invariant, Hck-encoded regions in these RRT-SH3 domains, only ligands with certain types of PxxP motifs could be efficiently targeted by this library. If necessary, such restrictions might be overcome by subjecting the library to further diversification by random mutagenesis. Alternatively, the library could be supplemented with sublibraries of similarly modified RRT-SH3 molecules constructed from 2 or 3 other divergent SH3 domains.

An obvious implication of the present study is that individual SH3-domains can have a capacity for mediating highly specific and strong interactions, and do not necessarily need the assistance of other binding domains to achieve this. One reason for the common assumption to the contrary could be that most studies on SH3-binding have involved only relatively short ligand peptides, which may not have been able to accommodate the RT-loop or other possible SH3 specificity determinants as efficiently as the corresponding native ligand proteins. On the other hand, it is also possible that it has rarely been beneficial in the evolution to develop SH3 domains with such extreme binding affinities as the artificial SH3 domains produced in this study, because

SH3 domains are typically utilized in signaling proteins involved in finely regulated and transient multiprotein complexes.

However, regardless of whether the affinities of most naturally occurring SH3-interactions are high or relatively low (say below or over  $K_D$  1  $\mu$ M), it is clear that artificial RRT-SH3 domains like the ones generated here against Nef, could be potent competitors for natural SH3 interactions. By such targeting of cellular SH3-ligand proteins, it could be possible to inhibit or otherwise regulate their functions and complexes with other proteins, or to control their subcellular localization by equipping the RRT-SH3 protein with an appropriate targeting signal. While the most apparent uses of such approaches are in basic research, in the future RRT-SH3 domains could have similar applications in gene therapy. Finally, structural analysis of the RRT-SH3/ligand complexes could also help in design of non-peptide compounds targeted against the SH3-binding interfaces of various disease-associated cellular proteins. Having now at hand a number of RRT-SH3 targeted with high-affinity against HIV-1 Nef, it will be interesting to test the above ideas in experiments aimed at elucidating the role of Nef in HIV replication cycle and AIDS, and in developing strategies for interfering with these functions.

669250 "SAGE" 052699

Attachment for a provisional application for a patent entitled:  
"Methods and materials for generating SH3 domains with tailored binding properties".  
Applicant: Saksela, Kalle Martti

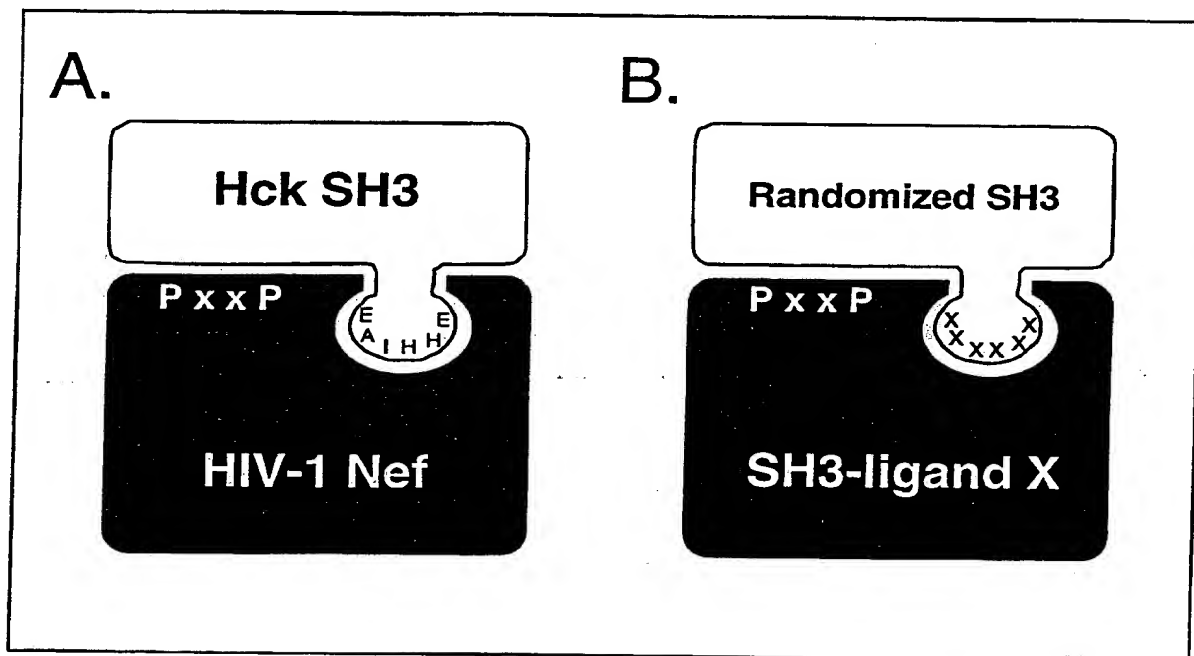


Figure 1

Attachment for a provisional application for a patent entitled:  
"Methods and materials for generating SH3 domains with tailored binding properties".  
Applicant: Sakselä, Kalle Martti

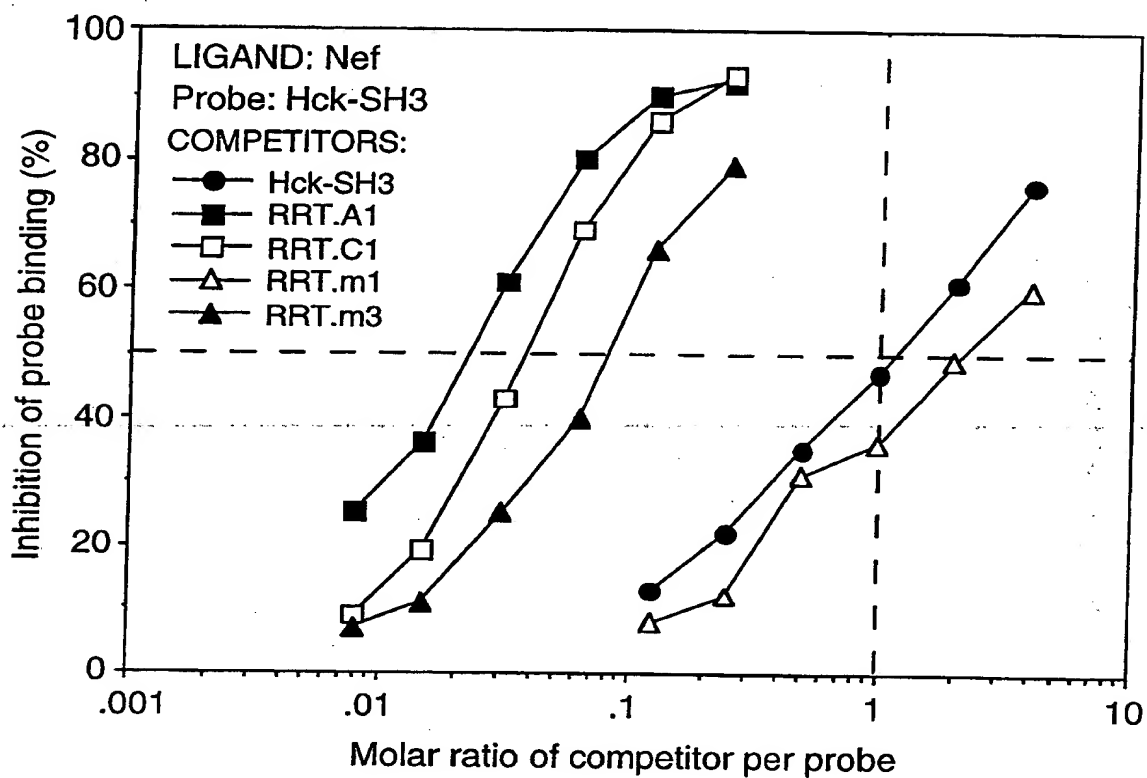


Figure 2

Attachment for a provisional application for a patent entitled:  
"Methods and materials for generating SH3 domains with tailored binding properties".  
Applicant: Saksela, Kalle Martti

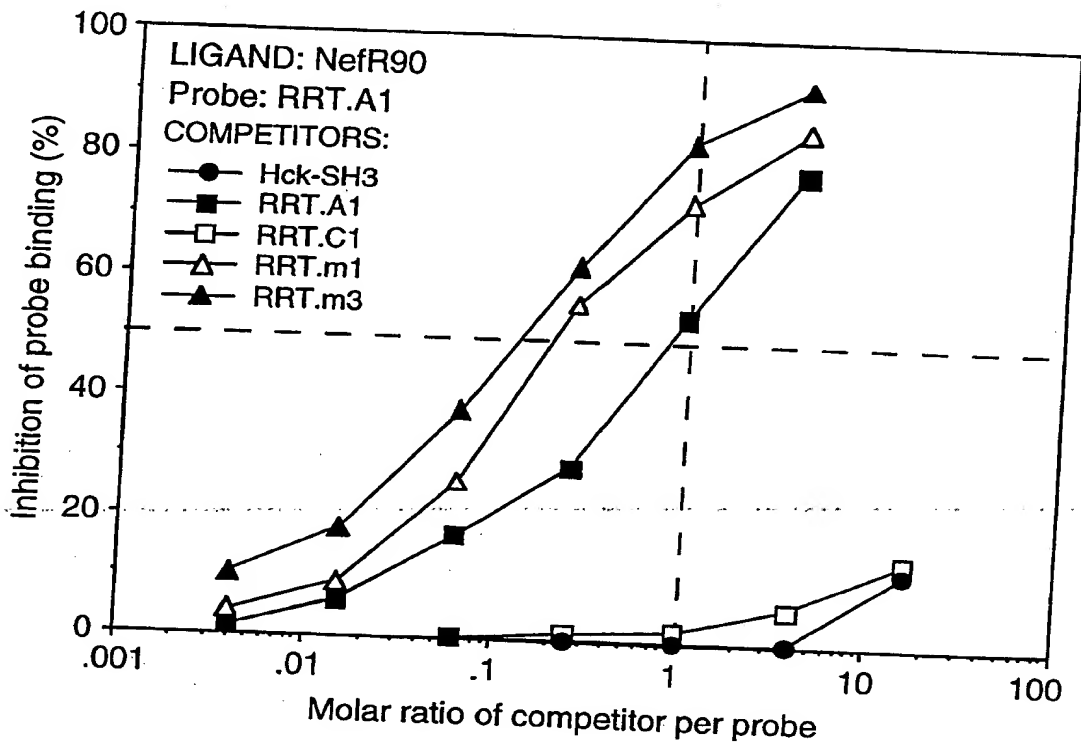


Figure 3